AMENDMENTS TO THE CLAIMS

The following list of claims will replace all prior versions and listings of claims in the application.

1. (Currently Amended) A method of determining the language of a textual passage, the method comprising the steps of:

(a) parsing said textual passage into ~~a plurality of~~ plural n-grams;

(b) comparing each of said plural n-grams with a plurality of databases, wherein each of said databases comprises ~~a list of~~ multiple n-grams associated with a specific language;

(c) determining an initial weight for each of said plural n-grams, per language, by calculating the frequency with which each of said plural n-grams appears in each of said databases and dividing said frequency by the total number of the multiple n-grams in said respective database;

(d) determining the number of said databases within which each of said plural n-grams appear;

(e) altering said initial weight for each of said plural n-grams by multiplying said initial weight with the inverse of said number of databases within which each of said plural n-grams appear to produce an altered weight for each of said plural n-grams;

~~(f) producing the weight of each language over the text passage by calculating, per language, the sum over each n-gram in the text passage of the products of the number of times that that n-gram appears in the text passage and the language-specific altered weight calculated in step (e) for that n-gram;~~

(f) multiplying, for each of the plural n-grams, a number of times that n-gram appears in the textual passage by the altered weight of that n-gram from step (e) to produce a product for each n-gram, per language, and summing those products to produce a language passage weight for each language for the textual passage;

(g) ~~sorting the list of per~~ determining, based upon a comparison of the language passage weights from step (f) ~~in decreasing order~~, ~~returning~~ the most likely language or languages for the ~~text~~ textual passage ~~as the first element (highest weight) in the list~~.

2. (Canceled) ~~The method of claim 1 wherein the step of determining an initial weight for each of said n-grams, per language, comprises the steps of calculating the frequency with which each of said n-grams appears in each of said databases and dividing said frequency by the total number of n-grams in said respective database.~~

3. (Currently Amended) The method of claim 1 wherein said plural n-grams have a size selected from the group consisting of bi-grams, tri-grams, and quad-grams.

4. (Currently Amended) The method of claim 1 wherein said plural n-grams are anchored n-grams.

5. (Currently Amended) The method of claim 1 wherein said plural n-grams are replacement-type n-grams.

6. (Currently Amended) The method of claim 1 wherein said plural n-grams are any combination of n-grams, including anchored n-grams and/or replacement-type n-grams, and/or n-grams of different lengths.

7. (Original) The method of claim 1 wherein said textual passage comprises 20 or more words.

8. (Original) The method of claim 1 wherein said textual passage comprises 40 or more words.

9. (Currently Amended) A method of determining the language of a textual passage, the method comprising the steps of:

(a) filtering ~~a plurality of~~ plural short words from a textual passage;

(b) comparing each of said plural short words against a plurality of databases, wherein each of said databases comprises ~~a list of~~ multiple short words associated with a different language;

(c) determining an initial weight for each of said plural short words, per language, by calculating the frequency with which each of said plural short words appears in each of said

databases and dividing said frequency by the total number of the multiple short words in said respective database;

(d) determining the number of said databases within which each of said plural short words appear;

(e) altering said initial weight for each of said plural short words by multiplying said initial weight with the inverse of said number of databases within which each of said plural short words appear to produce an altered weight for each of said plural short words;

~~producing the weight of each language over the text passage by calculating, per language, the sum over each short word in the text passage of the products of the number of times that that short word appears in the text passage and the language specific altered weight calculated in step (e) for that short word;~~

(f) multiplying, for each of the plural short words, a number of times that short word appears in the textual passage by the altered weight of that short word from step (e) to produce a product for each short word, per language, and summing those products to produce a language passage weight for each language for the textual passage;

(g) ~~sorting the list of per~~ determining, based upon a comparison of the language passage weights from step (f) ~~in decreasing order~~, ~~returning~~ the most likely language or languages for the ~~text~~ textual passage ~~as the first element (highest weight) in the list~~.

10. (Currently Amended) A method of determining the language of a textual passage, the method comprising the steps of:

(a) filtering ~~a plurality of~~ plural short words from a textual passage and parsing said textual passage into ~~a plurality of~~ plural n-grams;

(b) comparing each of said plural n-grams and said plural short words against a plurality of databases, wherein each of said databases comprises ~~a list of~~ multiple n-grams and short words associated with a different language;

(c) determining an initial weight for each of said plural n-grams and said plural short words, per language;

(d) determining the number of said databases within which each of said plural n-grams and said plural short words appear;

(e) altering said initial weight for each of said plural n-grams and said plural short words by multiplying said initial weight with the inverse of said number of databases within

which each of said plural n-grams and said plural short words appear to produce an altered weight for each of said plural n-grams and said plural short words;

~~producing the weight of each language over the text passage by calculating, per language, the sum over each short word and each n-gram in the text passage of the products of the number of times that that short word or n-gram appears in the text passage and the language-specific altered weight calculated in step (e) for that short word or n-gram;~~

(f) multiplying, for each of the plural n-grams and each of the plural short words, a number of times that n-gram or short word appears in the textual passage by the altered weight of that n-gram or short word from step (e) to produce a product for each n-gram and each short word, per language, and summing those products to produce a language passage weight for each language for the textual passage;

(g) ~~sorting the list of per~~ determining, based upon a comparison of the language passage weights from step (f) ~~in decreasing order~~, ~~returning~~ the most likely language or languages for the ~~text~~ textual passage ~~as the first element (highest weight) in the list~~.

11. (Currently Amended) A system for determining the language of a textual passage, ~~comprising:~~ said system comprising a central processing unit coupled to a memory system and a display, wherein said central processing unit operates according to a program retrieved from said memory system, wherein said program ~~includes the steps of;~~ is configured to cause the central processing unit to execute steps comprising:

(a) receiving a textual passage;

(b) parsing said textual passage into ~~a plurality of~~ plural n-grams;

(c) comparing each of said plural n-grams against a plurality of databases, wherein each of said databases comprises ~~a list of~~ multiple n-grams associated with a different language;

(d) assigning an initial weight to each of said plural n-grams, per language, by calculating the frequency with which each of said plural n-grams appears in each of said databases and dividing said frequency by the total number of n-grams in said respective database;

(e) calculating the number of said databases within which each of said plural n-grams appear;

(f) altering said initial weight assigned to each of said plural n-grams by multiplying said initial weight with the inverse of said number of databases within which each of said plural n-grams appear to produce an altered weight for each of said plural n-grams;

~~(g) producing the weight of each language over the text passage by calculating, per language, the sum over each n-gram in the text passage of the products of the number of times that that n-gram appears in the text passage and the language-specific altered weight calculated in step (f) for that n-gram;~~

(g) multiplying, for each of the plural n-grams, a number of times that n-gram appears in the textual passage by the altered weight of that n-gram from step (f) to produce a product for each n-gram, per language, and summing those products to produce a language passage weight for each language for the textual passage;

(h) ~~sorting the list of per~~ determining, based upon a comparison of the language passage weights from step (g) ~~in decreasing order~~, ~~returning~~ the most likely language or languages for the ~~text~~ textual passage ~~as the first element (highest weight) in the list~~.

12. (Original) The system of claim 11 further comprising a scanner and an optical character recognition device, wherein said scanner and said optical character recognition device are connected to said central processing unit, wherein said program receives a textual passage from a document scanned by said scanner.

13. (Original) The system of claim 11 wherein said program comprises a user interface that allows a user to enter said textual passage.

14. (Original) The system of claim 13 wherein said user interface is a graphical user interface.

15. (Original) The system of claim 13 wherein said user interface displays the identified language.

16. (Original) The system of claim 11 wherein said program comprises a user interface that allows a user to enter a Uniform Resource Locator that contains said textual passage.

17. (New) The method of claim 1, comprising determining the most likely language for the textual passage as that with the highest passage weight.

18. (New) The method of claim 9, comprising determining the most likely language for the textual passage as that with the highest passage weight.

19. (New) The method of claim 10, comprising determining the most likely language for the textual passage as that with the highest passage weight.

20. (New) The system of claim 11, wherein said program is configured to cause the central processing unit to determine the most likely language for the textual passage as that with the highest passage weight.